Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity[a)]

Nathaniel I. Durlach,[b)] Christine R. Mason, Barbara G. Shinn-Cunningham, Tanya L. Arbogast, H. Steven Colburn, and Gerald Kidd, Jr.
Hearing Research Center, Boston University, Boston, Massachusetts 02215

(Received 27 September 2002; revised 20 March 2003; accepted 8 April 2003)

Previous work has indicated that target-masker similarity, as well as stimulus uncertainty, influences the amount of informational masking that occurs in detection, discrimination, and recognition tasks. In each of five experiments reported in this paper, the detection threshold for a tonal target in random multitone maskers presented simultaneously with the target tone was measured for two conditions using the same set of five listeners. In one condition, the target was constructed to be "similar" (S) to the masker; in the other condition, it was constructed to be "dissimilar" (D) to the masker. The specific masker varied across experiments, but was constant for the two conditions. Target-masker similarity varied in dimensions such as duration, perceived location, direction of frequency glide, and spectro-temporal coherence. Group-mean results show large decreases in the amount of masking for the D condition relative to the S condition. In addition, individual differences (a hallmark of informational masking) are found to be much greater in the S condition than in the D condition. Furthermore, listener vulnerability to informational masking is found to be consistent to at least a moderate degree across experiments. © *2003 Acoustical Society of America.*
[DOI: 10.1121/1.1577562]

PACS numbers: 43.66.Dc, 43.66.Lj [MRL]

I. INTRODUCTION

There is considerable room for argument about how best to define informational masking or even whether the term "informational masking" is useful. Nevertheless, it is generally agreed that informational masking is distinct from energetic masking, where energetic masking is defined as the masking that results from competition between target and masker at the periphery of the auditory system (i.e., at the level of the basilar membrane or auditory nerve). Consistent with this characterization is the idea that informational masking reflects vulnerability of certain central portions of the auditory processing system (e.g., related to attentional phenomena). Furthermore, it has been amply demonstrated that substantial amounts of informational masking can be created through the introduction of uncertainty in the acoustic stimulus. In fact, some investigators have used the effects of uncertainty to define informational masking (e.g., Watson and Kelly, 1981; Neff, 1995; Oh and Lutfi, 2000). Although it has not been shown that stimulus uncertainty is either a necessary or a sufficient condition to produce nonenergetic masking, there is no doubt that stimulus uncertainty can produce large amounts of such masking under a wide variety of conditions. Further comments on some of these conceptual issues are available in Durlach *et al.* (2003a).

In this paper, we report the results of a series of detection experiments (involving tonal targets and random multitone maskers presented simultaneously with the target) designed to demonstrate that informational masking resulting from stimulus uncertainty can be substantially reduced by introducing target-masker dissimilarity. To the extent that this is in fact the case, one can conclude that an adequate theory of informational masking must take account of target-masker similarity as well as uncertainty. The experimental results reported in this paper add to previously reported informational masking experiments concerned with target-masker similarity not only by providing additional data, but by exploring additional dimensions of similarity and by employing the same set of listeners in all of the experiments. Inasmuch as intersubject variability is known to be extraordinarily large in informational masking experiments, comparing the results for individual listeners across several different experiments can provide important additional insight into the nature of informational masking.

II. BACKGROUND

Previous research on informational masking includes studies of both simultaneous and sequential masking for discrimination and recognition as well as for detection, for a wide variety of signal types ranging from simple tonal stimuli to running speech. The following comments focus on the empirical work on informational masking for simultaneous nonspeech stimuli in which the target is a fixed-frequency tone and the masker is a multitone complex whose frequency components are varied randomly from presentation to presentation or trial to trial (e.g., Neff and Green, 1987; Neff and Callaghan, 1987; Neff *et al.*, 1993; Neff, 1995; Neff and Dethlefs, 1995; Kidd *et al.*, 1994; Oh and Lutfi, 1998; Wright and Saberi, 1999; Richards *et al.*, 2002). The spacing of the masker components is restricted in such a way that relatively little masker energy occurs in the fre-

---

[a)]Portions of this work were presented at the Acoustical Society of America meeting in Pittsburgh, PA in June 2002 [Mason *et al.*, J. Acoust. Soc. Am. **111**, 2470(A) (2002)].
[b)]Electronic mail: durlach@mit.edu

quency region around the signal tone (called the ''protected region''). Because the amount of energetic masking decreases with the size of the protected region (cf. Neff *et al.*, 1993), many of the experiments designed to focus on informational masking use protected regions that are equal to or greater than the ''critical band'' around the given target frequency. Also, because in many cases both energetic and informational masking are expected to occur at least to some degree, attention is given to how much masking is energetic and how much is informational, and to how the two types of masking interact (e.g., Lutfi, 1990). The number of tonal components in the masker, and the frequency range of those components, as well as the extent to which the components are randomized in amplitude and frequency, varies with the experiment. Also, the relative effect of randomizing the spectrum of the masker between intervals and between trials in two-interval paradigms has been examined (cf. Neff and Green, 1987; Neff, 1995; Neff and Dethlefs, 1995; Wright and Saberi, 1999; Richards *et al.*, 2002).

The effect of randomizing the spectrum of the masker can be exceedingly large. However, the results of such experiments are strongly listener dependent. Whereas some listeners, occasionally referred to as ''holistic'' or ''synthetic'' listeners, evidence very large effects of the uncertainty in the multitone masker, other listeners, often referred to as ''analytic'' listeners, show hardly any effect at all (Espinoza-Varas and Watson, 1989; Neff and Dethlefs, 1995; Lutfi *et al.*, 2003). Moreover, it appears that the variation in the size of this effect arises primarily from variation in the masked threshold for the uncertain-masker case rather than for the certain-masker case (which is often broadband noise). Questions of current interest in this area include: To what extent does a listener's ability to resist informational masking vary with the experimental task? What other differences among listeners correlate with this ability? How much can this ability be enhanced by training? According to a recent study by Oxenham *et al.* (2003), there is a significant positive correlation between resistance to informational masking and musical training.

Despite the large amount of data on informational masking that has become available over the past few years, there have been only a few attempts to model informational masking. Currently there is no model that satisfactorily accounts for all of the empirical results, even when limited to the body of work on detecting a target tone in a simultaneous random multitone masker discussed above. The most extensive effort to date is the CoRE (component relative entropy) model proposed by Lutfi (1993). Oh and Lutfi (1998) have shown that the CoRE model, which uses the weighted outputs (mean levels and variances) of a set of peripheral filters in addition to a variable bandwidth ''attentional'' filter, can predict the variation in threshold with number of masker tones (as originally found by Neff and Green, 1987) with considerable accuracy. In other cases, however, such as the detection threshold for an inharmonic tone embedded in a randomized harmonic multitone masker, the model is less successful (Oh and Lutfi, 2000). In distinct but related efforts, both Wright and Saberi (1999) and Richards *et al.* (2002) have interpreted informational masking data in terms of channel-weighting analyses.

Apart from the modeling work noted above, which is focused primarily on uncertainty in the stimulus combined with channel weights, the main theoretical notions that have been proposed to help understand informational masking phenomena concern the perceptual grouping or segregation of target and masker (Leek *et al.*, 1991; Kidd *et al.*, 1994; Neff, 1995; Oh and Lutfi, 2000). At a crude intuitive level, informational masking occurs because the listener finds it difficult to focus attention on the target in the presence of a distracting or confusing masker. Although uncertainty is clearly relevant to this phenomenon, so is the extent to which the target ''sounds like'' the masker and is grouped with the masker. In the words of Leek *et al.* (1991, pp. 205–206), ''Informational masking is broadly defined as a degradation of auditory detection or discrimination of a signal embedded in a context of other similar sounds'' and ''A target that is sufficiently different from the surrounding tones along some acoustic dimension will be heard with increased precision.'' Thus, in addition to uncertainty, similarity, which is well known to be a factor in the extent to which auditory objects may be grouped into a single auditory image or segregated into separate images (e.g., Bregman, 1990), has also been considered as an important factor in informational masking.[1]

In the studies by Kidd *et al.* (1994) and Neff (1995), informational masking was reduced by decreasing the similarity between target and masker in a variety of dimensions including spectro-temporal pattern, duration, and interaural (i.e., spatial) relationship. In the Neff study (1995, p. 1910), the purpose was ''to increase the perceptual differences between the signal and the sinusoidal masker components and thus facilitate hearing out the signal from the tonal complex.'' In the Kidd *et al.* study (1994, p. 3475), the stimulus manipulations were chosen so that ''they produced the subjective impression that the signal and masker were perceptually segregated into different auditory 'objects' or 'images.' '' The results of both of these studies are considered along with our own results in Secs. IV and V. In the paper by Oh and Lutfi (2000) on harmonicity mentioned above, the authors state that ''large elevations in threshold are often attributed to the lack of any predictable structure in the masker that would allow listeners to segregate the single spectral component belonging to the signal from the collection of the components belonging to the masker'' (p. 706). Accordingly, they hypothesized that a harmonic masker should produce less masking when the target is not one of the harmonic components. Their results were consistent with this conjecture about the role of perceptual segregation in reducing informational masking (but could not be accounted for by the CoRE model).

In general, it seems clear that the amount of informational masking cannot be predicted solely on the basis of uncertainty (even when the computation of uncertainty goes beyond consideration of the uncertainty in the stimulus waveforms) and that target-masker similarity and the phenomena of grouping and segregation must also be considered. Although the extent to which the components of a complex acoustic stimulus are grouped and segregated into

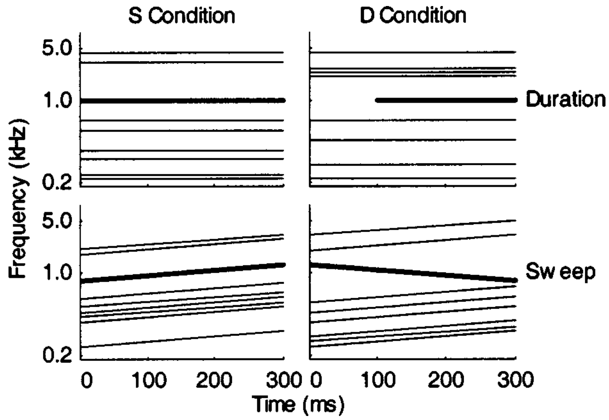

FIG. 1. Schematic illustration of the Duration experiment (top panels) and the Sweep experiment (bottom panels). All four panels show the case when the target is present (bolder lines). The two panels on the left illustrate the S condition and the two on the right the D condition. In the Duration experiment, the S and D conditions differ by target duration. In the Sweep experiment, the S and D conditions differ by the target sweep direction.

distinct images is difficult to quantify, there is some hope that eventually one or more metrics of target-masker similarity can prove useful in predicting the amount of informational masking that occurs. It should also be noted that target-masker similarity appears to be important in a wide range of complex auditory detection and recognition tasks. For example, there is substantial evidence that target-masker similarity plays a major role in speech reception tasks: informational masking tends to increase as the masker goes from noise to speech to same-sex talker to same talker (e.g., Freyman *et al.*, 1999, 2001; Brungart, 2001; Brungart *et al.*, 2001; Arbogast *et al.*, 2002).[2] Furthermore, a recent study by Kidd *et al.* (2002) provides support for the proposition that target-masker similarity affects informational masking for nonspeech pattern recognition. Finally, it should be noted that similarity is a well-known factor in the degree to which stimuli interfere with or mask each other in sequential as well as simultaneous masking [for extensive work on temporal patterns and sequential masking, see the work by Watson and his colleagues as exemplified in Watson *et al.* (1976), Watson and Kelly (1981), Watson (1987), and Espinoza-Varas and Watson (1989)] and in sensory channels other than audition [see, for example, Turvey (1973) for a consideration of pattern masking in vision].

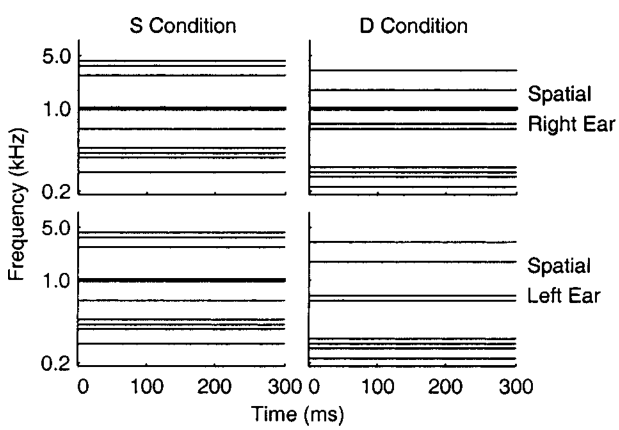

FIG. 2. Schematic illustration of the Spatial experiment. All panels show the case when the target is present (bolder lines). The top panels are the stimuli presented to the right ear and the bottom panels are the stimuli for the left ear. In the S condition (left panels), both the target and masker are presented diotically. In the D condition (right panels), the masker is presented diotically and the target monotically.

The purpose of the present study was to examine informational masking, and release from informational masking, for conditions in which target-masker similarity was varied while masker uncertainty remained unchanged. This work is thus a relatively direct extension of the previous investigations by Kidd *et al.* (1994), Neff (1995), and Oh and Lutfi (2000). In addition to testing new conditions in which target-masker similarity is varied, all of the listeners participated in all conditions of each experiment, thus allowing a determination of the consistency of listener performance across a diverse set of masking conditions. The expectation is that this data set will prove useful in future efforts to model informational masking that take into account the similarity and/or perceptual grouping and segregation of sounds.

III. EXPERIMENTS

A. Overview

Schematic diagrams of the experiments performed are shown in Figs. 1–3. In all cases, S is used to denote ‘‘target and masker similar’’ and D to denote ‘‘target and masker dissimilar.’’ The following paragraphs contain brief descriptions of the experiments performed (further details about these experiments are given in Sec. III B). In each case, a multitone masker with a high degree of frequency uncertainty is used. The distinction between the S and D conditions is made by changes to the target only and therefore involves no change in the masker uncertainty. In each case, it is intuitively obvious that the masked threshold in the D condition should be lower than in the S condition, despite the fact that such a result cannot reasonably be predicted on the basis of either stimulus energy or stimulus uncertainty (e.g., only in the fifth experiment is there any decrease in stimulus uncertainty in going from the S condition to the D condition, and even in that case it seems very unlikely that uncertainty rather than target-masker similarity is the relevant issue).

Although all the experiments performed were alike in the sense that the uncertainty in the stimulus consisted of frequency uncertainty in the masker, they differed with respect to the parameters used to manipulate the degree of similarity between target and masker. The first experiment made use of duration, the second of direction of frequency sweep, the third of interaural parameters influencing spatial perceptual characteristics, and the fourth and fifth of parameters influencing grouping and streaming perceptual characteristics. Taken all together, it is believed that the array of stimulus parameters (and the subjective counterparts of these parameters) manipulated constitutes a sample that is sufficiently broad to enable one to draw general conclusions about the interaction of similarity and uncertainty with reasonable safety.

1. The shortened-target-duration experiment (Duration)

As shown in the top two panels of Fig. 1, the S and D conditions differ by target duration only: the target in the D

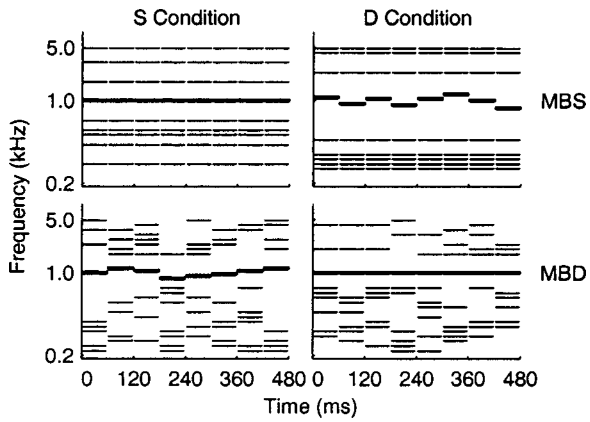

FIG. 3. Schematic illustration of the MBS and MBD experiments. All graphs show the case when the target is present (bolder lines). The top two panels show the MBS paradigm. In this case, the masker was always held fixed from burst to burst. In the S condition (top left), the target was also held fixed; in the D condition (top right), it was jittered in frequency from burst to burst. The bottom two panels show the MBD paradigm. In this case the masker was always jittered. In the S condition (bottom left), the target was also jittered; in the D condition (bottom right), the target was held fixed.

condition is shorter than the target in the S condition (i.e., the target is turned on after the onset of the masker). This experiment is similar to one performed by Neff (1995).

2. The reversed-frequency-sweep experiment (Sweep)

As shown in the bottom two panels of Fig. 1, the masker tones are all upward frequency glides. In the S condition, the target is a glide with the same extent and direction as the masker components. In the D condition, the target glide is in the opposite direction from the masker components.

3. The separate-spatial-channels experiment (Spatial)

Figure 2 illustrates the third experiment. The S condition consists of a diotic multitone masker with a diotic tonal target. In the D condition, the target is simply removed from one ear so that the masker is diotic and the target is monotic. This experiment is similar to a condition in the Kidd *et al.* (1994) paper and is similar in intent to a binaural condition in Neff (1995).

4. The jittered-target-frequency experiment (MBS)

The top two panels of Fig. 3 illustrate the fourth experiment. In both the S and D conditions, the masker and target consist of multiple-burst stimuli. Whereas the frequencies of the masker components are always held fixed from burst to burst [referred to as multiple-bursts-same, or MBS, as in Kidd *et al.* (1994)], the frequency of the target tone is fixed in the S condition but jittered in the D condition.

5. The constant-target-frequency experiment (MBD)

The bottom two panels of Fig. 3 illustrate the fifth experiment. As in the fourth experiment, multiple-burst stimuli are employed in both the S and D conditions. However, in this experiment, whereas the frequencies of the masker are now always randomized from burst to burst [referred to as multiple-bursts-different, or MBD, as in Kidd *et al.* (1994)], the frequency of the target tone is now jittered in the S condition but held fixed in the D condition.

B. Methods

1. Listeners

Five university students between the ages of 20 and 24 (three male undergraduates and two female graduate students) participated in all experiments. All five listeners, denoted L1–L5, had participated in previous experiments in our laboratory but were selected solely on the basis of availability. They were paid for their participation and completed the experiments in five 2-h sessions (with breaks) over the course of 2 weeks.

2. General methods and procedures

The stimuli were generated at a 20-kHz sampling rate and low-pass filtered at 7.5 kHz. All masker bursts consisted of eight tones that were chosen randomly on a logarithmic frequency scale from the range 200–5000 Hz on every presentation, excluding the subregion 800–1250 Hz. The target was always contained within this protected subregion. Sounds were presented to listeners through matched TDH50 headphones while seated in individual sound attenuating rooms. Unless stated otherwise, the target was a 1000-Hz tone, the stimulus was presented monaurally to the right ear, and only one stimulus burst occurred in each interval of a trial. All experiments used a 2I, 2AFC two-down and one-up adaptive procedure with a fixed masker level of 60 dB per masker component (approximately 69 dB overall level) and an adaptive target level. Also, all experiments employed correct-answer, trial-by-trial, feedback.

The sessions began with several adaptive runs to estimate the unmasked target thresholds as well as to familiarize the listener with the targets. Next, masked thresholds were obtained alternating between the two conditions S and D after every two adaptive runs. A total of eight adaptive tracks, with a minimum of 50 trials and 9 reversals each, were obtained for every condition. Each adaptive track began with a step size of 4 dB that was changed to 2 dB after the third reversal. An even number of reversals, beginning with the fourth or fifth, were averaged to obtain one threshold estimate. To reduce learning effects, only the last 6 of the 8 threshold estimates were used in the final data analysis.

3. Specific stimuli

In the Duration experiment (Fig. 1, top), the duration of the eight-component masker was 300 ms (including 20 ms cosine-squared ramps for both onset and offset). In the S case, the target had exactly the same temporal characteristics as the masker. In the D case, the target began 100 ms later than the masker but retained the synchronous offset; thus its duration was only 200 ms (including its 20 ms cosine-squared ramps).

In the Sweep experiment (Fig. 1, bottom), the eight components of the masker had random starting frequencies (as in the above experiment), but instead of remaining constant they were rising frequency glides. The frequency of each component increased by a factor of 1.49 over the 300-ms duration of the masker. In order to maintain the 5000-Hz upper limit on the frequencies present in the masker, the highest possible starting frequency of any masker component was 3356 Hz (5000 Hz/1.49). In the S case, the target was an upward glide from 820 to 1220 Hz. In the D case, it was a downward glide covering the same frequency range. In all cases, each component had a duration of 300 ms including 20 ms cosine squared onsets and offsets.

In the Spatial experiment (Fig. 2), the masker was the same as that used in the Duration experiment, except that it was presented diotically rather than monotically. The target tone was presented synchronously with the masker either diotically (the S case) or monotically (the D case).

In the MBS experiment (Fig. 3, top), the masker and the target consisted of eight contiguous 60-ms bursts with cosine-squared onset and offset ramps of 10 ms and a total duration of 480 ms. The frequency of each component of the masker was always held constant throughout the stimulus presentation (i.e., from burst to burst but not interval to interval). In the S condition, the frequency of the target was also held constant from burst to burst; in the D condition, however, the target frequency was randomly jittered from burst to burst over the range 820–1220 Hz.

In the MBD experiment (Fig. 3, bottom), the bursts have the same temporal characteristics as in the MBS case. However, in contrast to the MBS case, the frequencies of the masker were always randomized burst to burst in the range 200–5000 Hz (excluding the protected region). Thus, in this experiment the S condition employed a jittered-frequency target (in the range 820–1220 Hz) and the D condition employed a constant-frequency target.

IV. RESULTS

Possible learning effects were checked by examining the slopes of the threshold versus repetition functions. Averaged over the listeners and experiments, these slopes were relatively shallow (-0.38 dB/repetition for the S condition and -0.15 dB/repetition for the D condition). However, there was considerable variation in these slopes (the standard deviations were 2.2 and 1.8 dB/repetition, respectively). In general, these data are not adequate to study learning effects. Although for many listeners and many conditions, the learning observed over the last six repetitions (the measurements used in subsequent analyses) was relatively minor, one certainly cannot claim that asymptotic performance was approached. As pointed out previously, the issue of training in informational masking is an important one and will require substantial future work.

The average amount of masking for each experiment, displayed in bar graph form, is shown in Figs. 4 and 5. Whereas Fig. 4 shows the results averaged over listeners, Fig. 5 shows the results for the individual listeners (L1–L5). In both figures, the black bars show results for the S condition (target and masker similar) and the white bars for the D condition (target and masker dissimilar). In all cases, the amount of masking was obtained by subtracting each individual's unmasked target threshold from their masked threshold. Unmasked target thresholds ranged from -5 to 19 dB SPL for the various listeners and targets. The error bars in Fig. 4 give the standard error over the five listeners whereas the error bars in Fig. 5 give the standard deviation over the six final adaptive runs for each listener. (Standard error was chosen for Fig. 4 and standard deviation for Fig. 5 for visual display purposes. Standard deviations for Fig. 4 are obtained by multiplying the results shown by $\sqrt{5}$. Standard errors for Fig. 5 are obtained by dividing the results shown by $\sqrt{6}$.) On average, the standard deviations across the six repetitions are 7.1 dB for the S condition and 5.4 dB for the D condition. The most striking result seen in Fig. 4 is that in all five

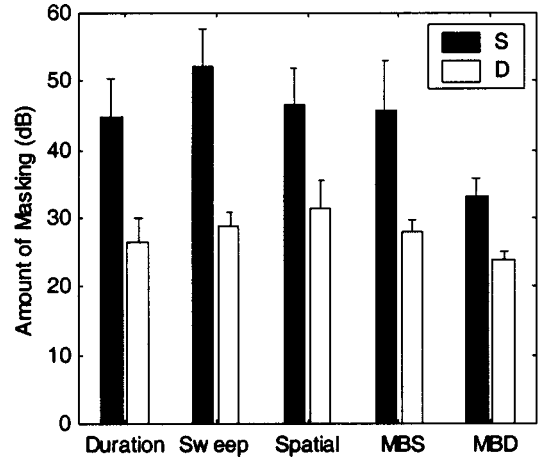

FIG. 4. Amount of masking (masked threshold minus unmasked threshold) for the five experiments and two target conditions S and D averaged over the five listeners. The results for the S condition are shown by black bars; the results for the D condition by white bars. The variation among listeners is indicated by the error bars, which show the standard error of the mean.

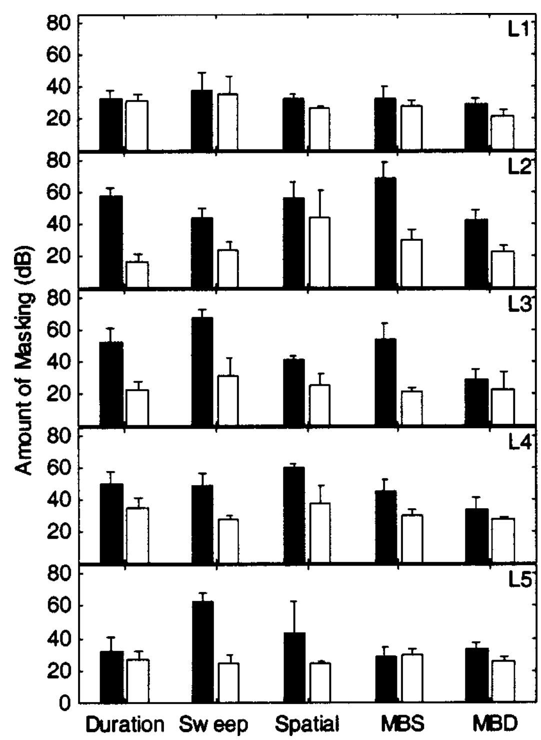

FIG. 5. The same as Fig. 4, except the results are plotted for each of the five listeners L1–L5 in separate panels. Here, the error bars give the standard deviation about the mean over the six adaptive runs used to estimate the threshold value.

TABLE I. This table shows the means (and standard deviations) for S, D, and S−D across experiments for each listener. The last column gives the means (and standard deviations) across listeners.

	L1	L2	L3	L4	L5	Mean (sd)
S	32.6(2.9)	53.9(10.8)	49.0(14.5)	47.3(9.5)	39.8(13.7)	44.5(8.4)
D	28.3(5.2)	27.6(10.6)	24.8(4.2)	31.5(4.6)	26.7(2.3)	27.8(2.5)
S−D	4.4(2.7)	26.3(12.9)	24.2(12.5)	15.8(6.4)	13.2(15.4)	16.8(8.8)

experiments there is considerably more masking for the S conditions than the D conditions. Given the specific target-masker parameters used and these specific five listeners, the most masking (average of approximately 52 dB) was obtained in the S condition for the Sweep experiment. The least masking for an S condition was obtained in the MBD experiment (approximately 33 dB). All of the D conditions produced less masking than their S counterparts and the order was somewhat different, with the Spatial experiment producing the most masking (approximately 32 dB) and the MBD experiment the least (an average of 24 dB). Despite these results and their statistical significance (discussed below), differences in the actual amount of masking obtained in the various experiments, the relative amount of masking across experiments, or the absolute size of the S−D difference will not be emphasized because of the dependence of these quantities on various arbitrary choices made in the design of the experiments (also discussed below).

Because listener differences are an important consideration (and we certainly do not assume that one can generalize the results obtained on five listeners to the population as a whole), a three-way ANOVA with listener as a factor was performed. The individual threshold estimates were used as the error term. All main effects and interactions were significant. Specifically, the three main factors of listener $[F(4,20)=9.97,\ p=0.0001]$, experiment $[F(4,20)=20.82,\ p<0.0001]$, and target-masker similarity $[F(1,5)=2951.3,\ p<0.0001]$ were all highly significant, indicating that each is an important determinant in the amount of masking produced. Perhaps more interesting are the interaction terms. The significant interaction of listener by experiment $[F(16,80)=10.27,\ p<0.0001]$ indicates that the difference across experiments depends on listener. The listener by target-masker similarity interaction $[F(4,20)=19.21,\ p<0.0001]$ indicates that the S vs D distinction also depends on listener, and the interaction of target-masker similarity and experiment $[F(4,20)=6.47,\ p=0.0017]$ indicates that the S vs D distinction also depends on experiment. The three-way interaction of listener by experiment by target-masker similarity $[F(16,80)=7.44,\ p<0.0001]$ indicates that the S vs D distinction depends on both listener and experiment. These points are elaborated below in connection with an examination of Figs. 4 and 5.

Looking first at Fig. 4, one sees that the average results across listeners clearly show the highly significant effect of target-masker similarity, as expected. The significant effect of experiment is also seen in that the experiments tend to produce different amounts of masking, with this difference appearing greater in the S condition than in the D condition. The interaction of experiment and target-masker similarity can be seen in that the effect of experiment is different for the S and D conditions. It is important to stress here, however, that a comparison of the amount of masking across experiments, or of the release from masking across experiments, calculated as S minus D (S−D), has limited meaning because these values could easily be changed by constructing the S conditions or the D conditions (or both conditions) differently. For example, of all the experiments performed, the Sweep experiment produces the most masking in the S condition and the greatest release from masking in going from the S condition to the D condition. However, the results for the Sweep experiment could obviously be radically changed merely by altering the slope of the sweep. Similarly, the results for the Duration experiment could be changed merely by altering the amount by which the target duration was shortened in the D condition. In general, no effort was made to equate (across experiments) either the effectiveness of the various maskers or the strength of the parameters used to produce the release from masking in the D conditions. Thus, in the remainder of this paper, comparisons of the amount of masking across experiments or the amount of release from masking across experiments will not be emphasized. Apart from the main result that in all experiments there is substantial release from masking in going from the S condition to the D condition, what is most interesting to explore in the data is how performance varied as a function of the listener, i.e., the results shown in Fig. 5.

Among the points to be noted when examining Fig. 5 are the following. First, as was implied in the average results, the D condition reduces the amount of masking relative to that obtained in the S condition in essentially all cases. Of the 25 comparisons between S and D shown in Fig. 5, 24 indicate a higher threshold for S than for D (the only exception is in the MBS experiment for L5). The mean values for S, D, and S−D (as well as the standard deviations) across experiments for each of the listeners (L1–L5) are shown in Table I along with the average and standard deviation of these mean values over the listeners. These results show quite clearly that the effect of target-masker similarity depends on listener (another of the significant two-way interactions in the ANOVA).

Second, not only is the amount of masking for D less than for S, but the intersubject variation in the amount of masking for D is less than for S. Although this result can be seen in Fig. 5, it is most clearly evident in the last column of Table I. The standard deviation across listeners for D is less than a third of the value for S (2.5 dB vs 8.4 dB).

Third, looking across both listeners and experiments in Fig. 5, one sees very large variations in the size of the S−D difference and in the amounts of masking for S and for D. The fact that the amount of masking in each experiment depends on the listener illustrates the significance of the two-way interaction of experiment and listener. Whereas, in some

TABLE II. This table shows the values of the quantity DP, which is a d' metric for measuring the difference between the results for the conditions S and D (taken from the values in Fig. 5), as well as the mean and standard deviation of DP across listener (last row) and experiment (last column). Bolded values are those differences that would be significant at the 0.05 level using a standard *t*-test given the Bonferroni correction (the critical DP value is 2.79, see text).

	Duration	Sweep	Spatial	MBS	MBD	Mean (sd)
L1	0.32	0.19	**2.89**	0.89	2.25	1.31 (1.2)
L2	**10.12**	**3.94**	0.96	**4.72**	**4.03**	4.76 (3.3)
L3	**4.64**	**4.82**	**3.05**	**5.10**	0.84	3.69 (1.8)
L4	2.27	**4.30**	**3.09**	**3.09**	1.22	2.79 (1.1)
L5	0.72	**8.04**	1.53	−0.33	1.92	2.38 (3.3)
Mean (sd)	3.61 (4.0)	4.26 (2.8)	2.31 (1.0)	2.69 (2.4)	2.05 (1.2)	

cases, the difference S−D is negligible, in other cases it is nearly 40 dB. The apparent dependence of this difference on both the experiment and the listener clearly reflects the statistical significance of the three-way interaction of these factors mentioned previously.

Fourth, it is evident in Fig. 5 that the variation among the error bars is extremely large (they vary by more than a factor of 10). In order to take these error bars into account when considering the difference S−D, a further analysis was performed. Specifically, in order to evaluate the S−D difference quantitatively, the d' metric DP, where

$$\mathrm{DP}=\frac{\overline{M}_S-\overline{M}_D}{\sqrt{(\sigma_S^2+\sigma_D^2)/2}}, \quad (1)$$

was computed for each experiment and listener. In this expression, $\overline{M}_S$ and $\overline{M}_D$ denote the means and σ_S and σ_D denote the standard deviations for the S and D cases, respectively. The results of these computations are shown in Table II. The bold values in this table indicate the conditions that are statistically significant at the 0.05 level using a standard *t*-test and the Bonferroni correction. The critical value of t for df=10, $p=0.05$ (corrected to $p=0.05/25=0.002$) is 4.144, which corresponds to a critical DP of 2.79 (in these calculations DP is $t/1.58$). According to this conservative analysis, roughly half of the DP values are significant. Whereas L2–L4 have three or four significant differences, L1 and L5 have only one significant difference. Note also that whereas four of the five listeners show significant differences in the sweep experiment, only one listener has a significant difference in the MBD experiment. These same conclusions can be drawn by looking across the panels of Fig. 5 at each listener or down the columns of the figure at each experiment. This dependence of the target-masker similarity effect on listener, as well as the dependence of the target-masker similarity effect on experiment, was confirmed by the significance of both of these two-way interactions in the ANOVA results. Again, however, the variation of the results across experiments must be interpreted with extreme caution because of the lack of a natural metric for equating the various stimulus alterations used to transform the S condition to the D condition (the values of DP would undoubtedly change if different magnitudes for the parameter manipulations were used to create the D conditions). The statistically significant two-way interaction of experiment and target-masker similarity would not necessarily remain significant if each D manipulation were somehow constructed to produce reductions in masking that were more equal across experiments.

Fifth, and finally, the results shown in Fig. 5 indicate an intermediate level of consistency of individual subjects across experiments (normalized to the level of performance of the average subject across experiments in order to factor out the arbitrary aspect of the interexperiment comparisons). For example, on the side of consistency, note that whereas L1 has relatively small S−D differences (primarily because of relatively low values for S), L2–L4 tend to have relatively large differences (primarily because of relatively high values of S). On the side of inconsistency, however, note how the results for L5 are like those for L1 in the Duration and MBD experiments, but not in the Sweep and Spatial experiments. A much fuller and more quantitative analysis of listener consistency across experiments and of differences across listeners, a significant main effect in the ANOVA results, is presented in the Appendix. According to the results obtained in that analysis, the S condition is distinguishable from the D condition not only by the larger amount of masking and the larger amount of intersubject variation, but also by an increased tendency for knowledge of listener identity to improve predictive accuracy across experiments. More specifically, the rms deviation between the measured amounts of masking and the amounts predicted by a simple linear model that takes account of listener identity was calculated for both S and D conditions. To test the model, the obtained rms deviation was compared to the probability distribution of rms deviations that would occur by chance (i.e., by ignoring listener identity). For the S condition, the obtained rms deviation (or smaller) would occur only 3% of the time by chance, indicating a substantial degree of listener consistency across experiments. In contrast for the D condition an rms deviation less than or equal to that obtained by including listener identity would occur in 52% of the cases without this knowledge.

V. COMPARISONS WITH PREVIOUS DATA

As mentioned in Sec. II, some of these experiments are closely related to experiments reported by Kidd *et al.* (1994) and Neff (1995). Precise quantitative comparisons with these previous experiments cannot be made because, aside from differences in the set of listeners employed, there are substantial differences in the details of the experiments. For example, in the Neff study, unlike our study, a target cue was always presented prior to each trial. Similarly, in the Kidd *et al.* multiple-burst experiments, unlike our multiple-burst experiments, the S condition was transformed into the D condition by altering the masker rather than the target. In addition, in neither study was the set of listeners held fixed across the experiments (thus preventing comparisons among studies of listener consistency across experiments). Nevertheless, to the extent that comparisons can be made across studies, the results appear relatively consistent. For example, using both a single-burst paradigm and a four-burst paradigm, for an eight-component masker in both a spatial experiment and a frequency-jitter experiment, Kidd *et al.* (1994) found large differences among listeners, with masking release in the range 0–40 dB (with an average of roughly 15 dB). Similarly, these same investigators found substantial release from informational masking (on the order of 20 dB) when the target was presented only during alternate bursts of the masker, despite the decrease in target energy in the alternate burst condition. In the Neff (1995) study, signal types, temporal factors, and spatial configuration were studied as a function of number of masker components. Even with the presence of a target cue immediately prior to each trial in all experiments (and trial-by-trial correct-answer feedback), substantial informational masking was obtained. Furthermore, and as expected, there was a substantial decrease in masking for most cases in which target-masker dissimilarity was introduced (again, differences among listeners were substantial). Relative to the baseline condition of a pure-tone target, AM (amplitude-modulated) targets, QFM (quasi-frequency-modulated) targets, and NBN (narrow-band noise) targets all showed decreased informational masking (although the QFM targets were least effective for this purpose). Most closely related to the experiments reported in this paper are the experiments with ten masker components in which the dissimilarity was created by shortening the signal duration or using different spatial configurations for the target and masker. The effect of making the target duration one-half the masker duration varied between 5 and 25 dB for the four listeners tested. Neff concluded that the duration manipulations were the most effective and most consistent at reducing the masking caused by masker frequency uncertainty. In that same study, the change in threshold in going from the monaural (or diotic) presentation for both target and masker to the case in which the target was presented interaurally out of phase varied over the range of 10–20 dB. As expected, a ‘‘cross-ear’’ condition in which the masker was in the ear opposite to the target produced the most release from masking although one listener required substantial practice before performance improved. In the spatial experiment by Kidd *et al.* (1994), for the case of 4 or 8 masker components, mean thresholds improved by 12–17 dB in going from the monotic to dichotic presentation (signal to one ear, masker to both ears in phase) for both the MBS and MBD (four-burst) conditions, and in going from the MBS to the MBD presentation for both the monotic and dichotic conditions.

For very rough comparisons with these previous data, the results of our experiments can be summarized as showing the following release-from-masking ranges and means (over the five listeners): a range of 6 to 22 dB and mean of 15 dB for Spatial; a range of 1 to 41 dB and mean of 18 dB for Duration; a range of −2 to 38 dB and mean of 18 dB for MBS; and a range 6 to 19 dB and mean of 9 dB for MBD. Because in both the experiments by Kidd *et al.* (1994) and Neff (1995) the listeners varied across experiments, the degree of listener consistency across experiments cannot be compared among the various sets of data.

VI. CONCLUDING REMARKS

The results reported in this paper, combined with those reported in the previous papers discussed above, clearly demonstrate that decreasing target-masker similarity (i.e., going from condition S to condition D) tends to reduce the masking effects of stimulus uncertainty. Also, as in previous experiments on informational masking, the intersubject variation is substantial. Furthermore, this variation appears much larger in the S condition than in the D condition. The amount of reduction (the threshold for S minus the threshold for D, S−D) depends both on the type of similarity change and on the listener. The results obtained in our experiments suggest that there is considerable structure in the matrix of thresholds for the ten different experimental conditions across the five different listeners. Specifically, in conditions where the amount of informational masking is reduced by decreasing target-masker similarity, individual differences in performance are relatively modest. In contrast, for conditions in which target and masker are similar, individual differences are large and the relative amounts of masking observed for a particular listener are moderately consistent across experiments that use different stimuli and methods of decreasing similarity. One cannot conclude, however, that individual differences are uniformly large in informational masking tasks and small in the reference tasks, or that a particular listener's vulnerability to informational masking is rigidly fixed across tasks. Such a conclusion would not only overstate the results obtained in this study, but other studies as well. For example, intersubject variability in informational masking tasks involving speech intelligibility seem somewhat reduced (Brungart, 2001; Arbogast *et al.*, 2002). Similarly, in some studies of informational masking, the intersubject differences in thresholds for the reference conditions (no uncertainty), as well as the informational masking conditions, appear quite large (e.g., Wright and Saberi, 1999; Durlach *et al.*, 2003b).

As indicated previously, it is difficult to draw conclusions about the relative potency in combating uncertainty of the different target-masker dissimilarity parameters introduced to convert the S condition to the D condition because of the arbitrary choice of the magnitudes of these parameters and the current lack of an independent metric to measure target-masker similarity. Despite this deficiency, all five of these particular experiments averaged across these particular five listeners produced substantial amounts of masking (between 33 and 52 dB) in the S conditions and large amounts of release from masking (9 to 23 dB) for the manipulations used to create the D conditions. Clearly, an important task for the future is to develop a target-masker similarity metric that can be applied to a wide variety of experimental situations.

A further issue that has not been addressed by the above experiments concerns the extent to which the observed releases from masking caused by the reductions of target-masker similarity in the various experiments would have occurred even if there had been no uncertainty in the masker. It has been implied implicitly by our use of the phrase ''combating uncertainty'' that if there were no uncertainty, there would be no nonenergetic masking for the decrease in target-masker similarity to combat. However, it is possible that even if the masker uncertainty had been totally eliminated, the decrease in target-masker similarity in going from condition S to condition D would have caused significant release from masking. In order to adequately explore this issue, it would have been necessary to measure thresholds for the S and D conditions in each experiment for all frozen exemplars of the random masker. To the extent that the results of this additional (massive) set of experiments showed a clear release from masking in going from condition S to condition D (and this release from masking were of sufficient magnitude to rule out explanations in terms of possible changes in energetic masking that might have occurred in some of the frozen cases in going from S to D), one would be forced either to define informational masking so as to include effects other than those associated with uncertainty OR to recognize that there exist types of nonenergetic masking other than informational masking. [More extended discussion of such definitional issues is available in Durlach *et al.* (2003a).]

It should also be noted that the data shown in Sec. IV of this paper cannot be compared to a quantitative theory of informational (or nonenergetic) masking because there is no such theory that now exists that takes quantitative account of both masker uncertainty and target-masker similarity. In order to develop such a theory, it will be necessary to define both uncertainty and similarity more adequately, determine improved methods for measuring these factors, and create a structure for properly integrating the effects of these factors.

It should further be noted that for such a theory to be truly successful, it will have to explain the very large individual differences observed as well as the effects of training (once these effects have been adequately documented empirically). Independent of whether the effects of training are generally large or small, and independent of the extent to which training tends to reduce the large individual differences observed, the study of training effects constitutes an essential step in the development of a serious theory. Although individual differences in susceptibility to informational masking are clearly of interest even if such differences can eventually be ''trained out,'' the way in which such differences should be modeled will obviously depend on the extent to which, and the manner in which, such ''training out'' can be achieved.

Finally, it should be noted that the stratagem in our research of focusing on simultaneous informational masking and temporarily ignoring sequential informational masking is not meant to imply that we believe that the latter area is unimportant or that an acceptable theory of informational masking can attend only to the simultaneous case. On the contrary, we believe that the results obtained on sequential masking (including results on individual differences and training effects) constitute a major building block in the search for an adequate theory. The area of sequential informational masking, and its relationship to simultaneous informational masking, will be considered in later papers.

ACKNOWLEDGMENTS

The authors thank Christine Carter and Sally Tressler for assistance with data collection and Dr. Robert Lutfi and Dr. Virginia Richards for useful discussions. In addition we are grateful to Dr. Marjorie Leek and two anonymous reviewers for their constructive suggestions. This work was supported by Grant Nos. DC00100 and DC04545 from NIH/NIDCD and Grant No. F49620-01-1-0005 from AFOSR.

APPENDIX: ANALYSIS OF LISTENER CONSISTENCY

One of the goals of the current set of experiments is to quantify the extent to which the large individual differences propagate across experiments. For example, if a particular listener has a very high threshold in the Duration experiment when target and masker have the same duration, does this listener also have a high a threshold in the Sweep experiment when both target and masker are swept in the same direction? How well does knowledge of individual performance in one task predict performance in another task, and is such individual knowledge more informative for S conditions (where individual differences are large and where informational masking is more important) than for D conditions? To what extent can individual listeners be characterized simply by determining their relative susceptibility to informational masking?

To begin to address these questions, we constructed a simple linear model of masking in which there is no interaction between the listener and the experiment (i.e., the effects of listener and experiment are completely separable[3]). We evaluated how well this model predicts the observed thresholds compared to predictions in which subject identity is ignored as well as predictions in which data were randomly permuted to destroy any listener consistency that might exist across experiments. For the S condition, the simple linear model that includes listener identity is shown to account for variability in the data beyond what one would expect by chance. However, for the D condition, the improvements in the model predictions that take into account listener identity are no more than would be expected by chance. These results, described below, suggest that listener differences are relatively consistent across different tasks when there is substantial informational masking, but not when there is little informational masking.

In the simple linear model, the amount of masking a particular listener exhibits in a particular experiment is assumed to be a sum of two factors: a factor specific to that experiment and a factor specific to that listener:

$$\mathrm{M(L,E)} - \overline{\mathrm{M(L,E)}}^{\mathrm{L,E}} = (\overline{\mathrm{M(L,E)}}^{\mathrm{L}} - \overline{\mathrm{M(L,E)}}^{\mathrm{L,E}}) + (\overline{\mathrm{M(L,E)}}^{\mathrm{E}} - \overline{\mathrm{M(L,E)}}^{\mathrm{L,E}}), \qquad \text{(A1)}$$

TABLE III. This table shows the rms deviation (RMS), correlation (r), and correlation squared (r^2) for both the S condition and the D condition for Eq. (A2) and Eq. (A3).

	S condition			D condition		
Model	RMS	r	r^2	RMS	r	r^2
Eq. (A2)	7.8	0.78	0.61	4.8	0.57	0.33
Eq. (A3)	10.8	0.49	0.24	5.3	0.43	0.18

where L denotes the listener and can assume any of the values L1,L2,...,L5; E denotes the experiment and can assume any of the values Duration, Sweep, Spatial, MBS, or MBD; M(L,E) denotes the amount of masking (in dB) for listener L and experiment E (as shown in Fig. 5); $\overline{\mathrm{M(L,E)}}^{\mathrm{L}}$ denotes the average of M(L,E) over L (the group mean profile shown in Fig. 4); $\overline{\mathrm{M(L,E)}}^{\mathrm{E}}$ denotes the average of M(L,E) over E (as reported in Table I); and $\overline{\mathrm{M(L,E)}}^{\mathrm{L,E}}$ denotes the average of M(L,E) over both L and E (the grand mean of all the data as reported in the last column of Table I).

Note that by collecting and rearranging terms, Eq. (A1) can be rewritten simply as

$$\mathrm{M(L,E)} = \overline{\mathrm{M(L,E)}}^{\mathrm{L}} + \overline{\mathrm{M(L,E)}}^{\mathrm{E}} - \overline{\mathrm{M(L,E)}}^{\mathrm{L,E}}. \tag{A2}$$

The relationship described by Eqs. (A1) and (A2) assumes that the results for listener L can be estimated by adding $\overline{\mathrm{M(L,E)}}^{\mathrm{E}}$ (a constant for each value of L) to the group-mean profile $\overline{\mathrm{M(L,E)}}^{\mathrm{L}}$, normalized by the overall group mean $\overline{\mathrm{M(L,E)}}^{\mathrm{L,E}}$. Note, furthermore, that equations (A1) and (A2) perfectly describe the data both when performance is the same for all listeners [because then $\mathrm{M(L,E)} = \overline{\mathrm{M(L,E)}}^{\mathrm{L}}$ and $\overline{\mathrm{M(L,E)}}^{\mathrm{E}} = \overline{\mathrm{M(L,E)}}^{\mathrm{L,E}}$ for all L and E] and when performance is the same for all experiments [because then $\mathrm{M(L,E)} = \overline{\mathrm{M(L,E)}}^{\mathrm{E}}$ and $\overline{\mathrm{M(L,E)}}^{\mathrm{L}} = \overline{\mathrm{M(L,E)}}^{\mathrm{L,E}}$ for all L and E].

In order to evaluate the extent to which Eq. (A2) represents the data for both the S and D conditions, the rms deviation between the predicted values of M(L,E) and the measured values of M(L,E) was computed (separately for S and D conditions). The results of this computation, included in Table III, show that the rms deviation for the S condition is 7.8 dB and the rms deviation for the D condition is 4.8 dB. If instead of using Eq. (A2) to estimate M(L,E), we used simply

$$\mathrm{M(L,E)} = \overline{\mathrm{M(L,E)}}^{\mathrm{L}}, \tag{A3}$$

i.e., we ignored subject differences and just used the group-mean profile to estimate M(L,E), then the rms deviations (also shown in Table III) would have been 10.8 dB for the S condition and 5.3 dB for the D condition. Although in an absolute sense, the rms deviation between data and predictions is larger in the S condition than in the D condition, subject differences account for a larger percentage of the variation in the S condition [(10.8−7.8) dB out of 10.8 dB or 28%] than in the D condition [(5.3−4.8) dB out of 5.3 dB or 8.6%].

An alternative way to compare Eqs. (A2) and (A3) is to calculate the correlations between the predicted and actual results in each case and determine the percentage of variation in the data for which the model accounts. These calculations (see Table III) show that 24% of the variance is accounted for in the S condition and 18% for the D condition when the mean alone is used [Eq. (A3)], but that these values increase to 61% in the S condition and 33% in the D condition when a listener-specific term is included in the predictions [Eq. (A2)]. Thus, incorporating knowledge of listener identity explains 37% more of the variance for the S condition, but only an additional 15% of the variance in the D condition (compared to using only knowledge of the experiment).

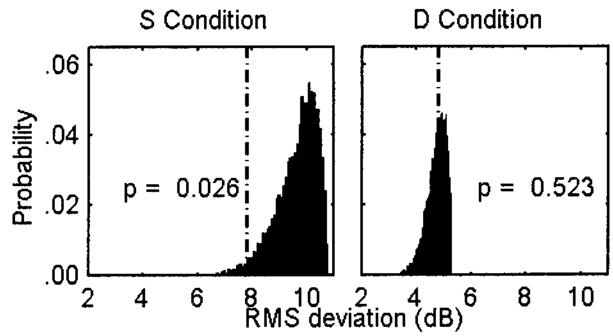

FIG. 6. Consistency of listeners across experiments. The figure shows the rms deviation between data and model for the both real listeners (dashed vertical line) and the pseudo listeners (probability density) for the S condition (left panel) and the D condition (right panel). See text for details.

While these analyses suggest that knowledge of listener identity improves prediction accuracy, Eq. (A2) has more degrees of freedom than Eq. (A3); thus it is not a "fair" comparison. Even if data points for each experiment are randomly assigned to "pseudo-listeners" [rather than grouping the data by actual listeners in calculating $\overline{\mathrm{M(L,E)}}^{\mathrm{E}}$], the rms deviation will always decrease using the more-complex model [Eq. (A2)] compared to the experiment-only model [Eq. (A3)]. In order to obtain better insight into this issue, a boot-strapping method was used to determine the extent to which, for the data points we were fitting, the observed improvements in the model predictions is more likely to arise by chance than from the actual listener-specific characteristics. More specifically, in order to assess whether these improvements are better than expected by chance, we determined how often random permutations of the measured data lead to better predictions than the predictions based on grouping the data by listener. In other words, predictions using Eq. (A2) were compared to the results obtained when the correspondence between listeners and the measured values of M(L,E) were randomized. In this analysis, we (a) constructed results for randomized pseudo listeners by randomizing the correspondence between L and M(L,E) (subject only to the constraint that the experiment E was held fixed in the randomization); (b) calculated the rms deviation for each such randomization in the same manner as described previously; (c) performed 10 000 such randomizations and rms-deviation computations; and (d) used these results to estimate the probability density of the rms deviations for these randomized pseudo listeners.

The results of this analysis are shown in Fig. 6. For each of the conditions S and D, the figure shows both the rms deviation obtained with the real listeners (represented by the dashed vertical lines) and the estimated probability density of the rms deviation for the pseudo listeners.[4] We also calculated the probability p of achieving an rms deviation with the pseudo listeners that is less than or equal to the actual rms deviation achieved with the real listeners (i.e., of achieving the actual rms deviation by chance). This quantity was calculated by estimating the area under the probability density functions to the left of the vertical line in Fig. 6. As shown, the probability of achieving this good a fit by chance is less than 3% in the S condition but is roughly 50% in the D condition. From this analysis, we conclude that compared to the D condition (in which the amount of informational masking is small), the results for the S condition (in which the amount of informational masking is large) show not only relatively large intersubject differences (see Table I), but also at least a modest degree of consistency across experiments. To what extent these results would continue to hold for other subjects and other experimental conditions is, of course, unknown.

Finally, it should be noted that although the relatively large rms deviations for the real listeners (7.8 dB for S and 4.8 dB for D) indicate that the model expressed by Eq. (A2) is far from perfect, even if the model were perfect these rms deviations would necessarily be substantially greater than zero because of the intrinsic noisiness of the M(L,E) measurements. In particular, note that if one averages the standard deviations of these measurements (displayed by the error bars in Fig. 5) across listeners and experiments, the results are 7.1 dB for condition S and 5.4 dB for condition D (corresponding to standard errors of 2.9 dB and 2.2 dB, respectively). With these numbers in mind, the rms deviations of 7.8 dB for S and 4.8 dB for D do not look so large.[5] In addition, more general models of listener consistency across experiments (nonlinear models, models incorporating parameters such as age, musical experience, etc.) may be more appropriate descriptions of how listener differences influence masking.[6] This simple linear representation was chosen because of its intuitive appeal and simplicity, not because we believe it represents the ''correct'' model of how listener differences propagate across experiments.

Overall, we interpret our results as confirming (1) the existence of large intersubject differences in susceptibility to informational masking and (2) substantial, but far from perfect, intrasubject consistency in susceptibility to informational masking across different types of informational masking experiments.

[1]It should be noted that we do not necessarily equate perceptual grouping (or fusion) with masking [see Bregman (1990) for a detailed discussion of this topic]. However, there is considerable evidence that when a target is perceptually grouped with an informational masker, the target is generally less detectable than when the stimuli are manipulated to promote segregation.

[2]The concept of masking due to perceptual similarity has been noted in the speech literature beginning at least with Egan *et al.* (1954) who distinguished between masking and ''confusion.'' Note also the close relationship between the terms ''perceptual masking'' [as used by Carhart *et al.* (1969)] and ''informational masking'' (as used by Freyman and colleagues and Brungart and colleagues in their recent speech-masking work).

[3]We realize that our data show some interaction between listener and experiment. This interaction is evident both when looking at Fig. 5 and when considering the results of the ANOVA. Nevertheless, it also appears that this interaction is sufficiently modest to warrant examination of listener consistency across experiments.

[4]Despite that fact that we refer to these randomizations as ''pseudo-listeners,'' it should be noted that they represent 10 000 variations of 25 values taken 5 at a time from the actual data obtained from only five listeners. This analysis is designed to examine whether the variations observed across the specific listeners used in this experiment are random or whether there are consistent differences among the results for these five particular subjects. This analysis is *not* equivalent to determining the variation that would be obtained if we actually observed 10 000 real listeners, but only considers whether the observed variation in the 25 data points we obtained is purely random or has some listener-specific structure.

[5]A rigorous evaluation of the model represented by Eq. (A2) would require modifying the computations underlying the results shown in Fig. 6 in such a way that the measurement noise is taken into account for both the real listeners and the pseudo listeners. We believe, however, that such a modification would add considerable complexity without providing much additional insight beyond that obtained with the simplified model used to derive the results shown in Fig. 6.

[6]A more general concept of listener consistency might require only that L and E be separable in the sense that functions H, f, and g exist such that M(L,E) can be well estimated by the equation $M(L,E)=H(f(L),g(E))$. In the model used in this paper [and ignoring the constant $\overline{M(L,E)}^{L,E}$], H merely adds f(L) and g(E), and f(L) and g(E) are chosen simply to equal $\overline{M(L,E)}^{E}$ and $\overline{M(L,E)}^{L}$, respectively. There is no guarantee, however, that this simple linear estimate is the best that could be found.

Arbogast, T. L., Mason, C. R., and Kidd, Jr., G. (**2002**). ''The effect of spatial separation on informational and energetic masking of speech,'' J. Acoust. Soc. Am. **112**, 2086–2098.

Bregman, A. S. (**1990**). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT, Cambridge, MA).

Brungart, D. S. (**2001**). ''Informational and energetic masking effects in the perception of two simultaneous talkers,'' J. Acoust. Soc. Am. **109**, 1101–1109.

Brungart, D. S., Simpson, B. D., Ericson, M. A., and Scott, K. R. (**2001**). ''Informational and energetic masking effects in the perception of multiple simultaneous talkers,'' J. Acoust. Soc. Am. **110**, 2527–2538.

Carhart, R., Tillman, T., and Greetis, R. (**1969**). ''Perceptual masking of spondees by combinations of talkers,'' J. Acoust. Soc. Am. **45**, 694–703.

Durlach, N. I., Mason, C. R., Kidd, Jr., G., Arbogast, T. L., Colburn, H. S., Shinn-Cunningham, B. (**2003a**). ''Note on informational masking,'' J. Acoust. Soc. Am. (in press).

Durlach, N. I., Mason, C. R., Arbogast, T. L., and Kidd, Jr., G. (**2003b**). ''Informational masking: Psychometric functions for 'frozen' and 'mixed' maskers,'' in Abstracts of Twenty-sixth Annual Mid-Winter Meeting of the Association for Research in Otolaryngology, Vol. 26, p. 224.

Egan, J. P., Carterette, E. C., and Thwing, E. J. (**1954**). ''Some factors affecting multi-channel listening,'' J. Acoust. Soc. Am. **26**, 774–782.

Espinoza-Varas, B., and Watson, C. S. (**1989**). ''Perception of complex auditory patterns by humans,'' in *Comparative Psychology of Audition: Perceiving Complex Sounds*, edited by R. J. Dooling and S. H. Hulse (Erlbaum, Hillsdale, NJ).

Freyman, R. L., Helfer, K. S., McCall, D. D., and Clifton, R. K. (**1999**). ''The role of perceived spatial separation in the unmasking off speech,'' J. Acoust. Soc. Am. **106**, 3578–3588.

Freyman, R. L., Balakrishnan, U., and Helfer, K. S. (**2001**). ''Spatial release from informational masking in speech recognition,'' J. Acoust. Soc. Am. **109**, 2112–2122.

Kidd, Jr., G., Mason, C. R., and Arbogast, T. L. (**2002**). ''Similarity, uncertainty and masking in the identification of nonspeech auditory patterns,'' J. Acoust. Soc. Am. **111**, 1367–1376.

Kidd, Jr., G., Mason, C. R., Deliwala, P. S., Woods, W. S., and Colburn, H. S. (**1994**). ''Reducing informational masking by sound segregation,'' J. Acoust. Soc. Am. **95**, 3475–3480.

Leek, M., Brown, M. E., and Dorman, M. F. (**1991**). ''Informational masking and auditory attention,'' Percept. Psychophys. **50**, 205–214.

Lutfi, R. A. (**1990**). ''How much masking is informational masking,'' J. Acoust. Soc. Am. **88**, 2607–2610.

Lutfi, R. A. (**1993**). ''A model of auditory pattern analysis based on component-relative-entropy,'' J. Acoust. Soc. Am. **94**, 748–758.

Lutfi, R. A., Kistler, D. J., Oh, E. L., Wightman, F. L., and Callahan, M. R. (**2003**). ''One factor underlies individual differences in auditory informational masking within and across age groups,'' Percept. Psychophys. **65**, 396–406.

Neff, D. L. (**1995**). ''Signal properties that reduce masking by simultaneous, random-frequency maskers,'' J. Acoust. Soc. Am. **98**, 1909–1920.

Neff, D. L., and Callaghan, B. P. (**1987**). *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, NJ), pp. 37–46.

Neff, D. L., and Green, D. M. (**1987**). ''Masking produced by spectral uncertainty with multicomponent maskers,'' Percept. Psychophys. **41**, 409–415.

Neff, D. L., and Dethlefs, T. M. (**1995**). ''Individual differences in simultaneous masking with random-frequency, multicomponent maskers,'' J. Acoust. Soc. Am. **98**, 125–134.

Neff, D. L., Dethlefs, T. M., and Jesteadt, W. (**1993**). ''Informational masking for multicomponent maskers with spectral gaps,'' J. Acoust. Soc. Am. **94**, 3112–3126.

Oh, E. L., and Lutfi, R. A. (**1998**). ''Nonmonotonicity of informational masking,'' J. Acoust. Soc. Am. **104**, 3489–3499.

Oh, E. L., and Lutfi, R. A. (**2000**). ''Effect of masker harmonicity on informational masking,'' J. Acoust. Soc. Am. **108**, 706–709.

Oxenham, A. J., Fligor, B., Mason, C. R., and Kidd, Jr., G. (**2003**). ''Informational masking and musical training,'' (in review).

Richards, V. M., Tang, Z., and Kidd, Jr., G. (**2002**). ''Informational masking with small set sizes,'' J. Acoust. Soc. Am. **111**, 1359–1366.

Turvey, M. T. (**1973**). ''On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli,'' Psychol. Rev. **80**, 1–52.

Watson, C. S. (**1987**). ''Uncertainty, informational masking and the capacity of immediateauditory memory,'' in *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, NJ), pp. 267–277.

Watson, C. S., and Kelly, W. J. (**1981**). ''The role of stimulus uncertainty in the discriminability of auditory patterns,'' in *Auditory and Visual Pattern Recognition*, edited by D. J. Getty and J. H. Howard, Jr. (Erlbaum, Hillsdale, NJ), pp. 37–59.

Watson, C. S., Kelly, W. J., and Wroton, H. W. (**1976**). ''Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty,'' J. Acoust. Soc. Am. **60**, 1176–1186.

Wright, B. A., and Saberi, K. (**1999**). ''Strategies used to detect auditory signals in small sets of random maskers,'' J. Acoust. Soc. Am. **105**, 1765–1775.